NOAA FISHERIES

Southeast Fisheries Science Center

Theme 1: Science and technical approaches

Suitability of the stock assessment models employed, taking into account constraints imposed by availability of data

Clay E. Porch (and many contributors)
Southeast Fisheries Science Center
July 2014

Overview

Stock assessments apply mathematical and statistical models to data collected from living resources and their associated fisheries to provide scientific advice on the status of those resources and the possible effects of management measures.

- Models for estimating stock status from data
- Models for generating scientific advice on management measures (e.g., ABC)

Models for estimating stock status from data

"Counting fish is just like counting trees — except that they are invisible and keep moving." John Sheperd

Preprocessing data

Incorporation of data in the synoptic assessment model

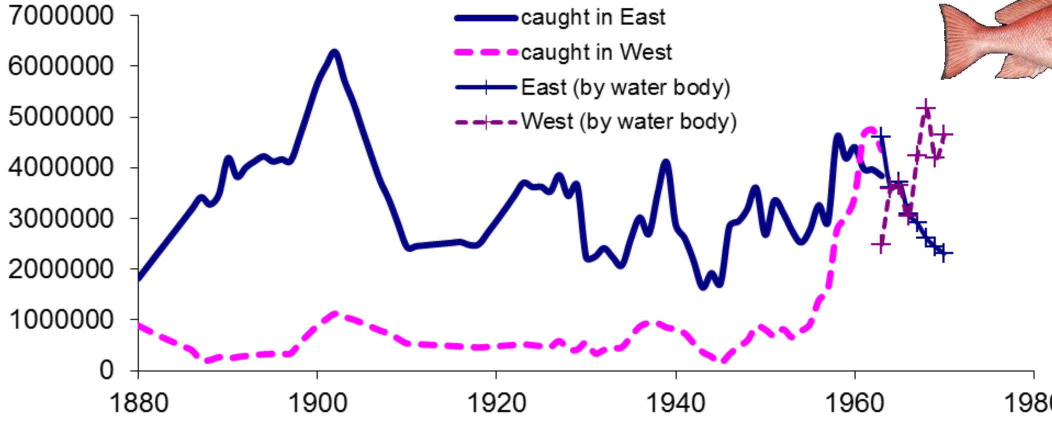
Structural assumptions in the assessment model

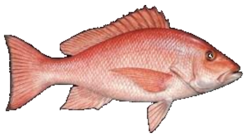
Model diagnostics

Landings

Reconstructing landings history prior to the 1980s

(before ALS and MRFSS/MRIP)





Pros: -Many fisheries already heavily exploited by 1980s (lack of contrast)

-Reduce illogical outcomes (population in overfished status, but estimated MSY > historical landings)

-Stabilize estimation (esp. if can assume unfished starting condition)

Cons: -Time consuming and difficult to defend for many species

(e.g, unknown, but large landings of red grouper by Cuba pre 1970s)

-Little other information available prior to 1980's

-Recreational particularly challenging

Landings

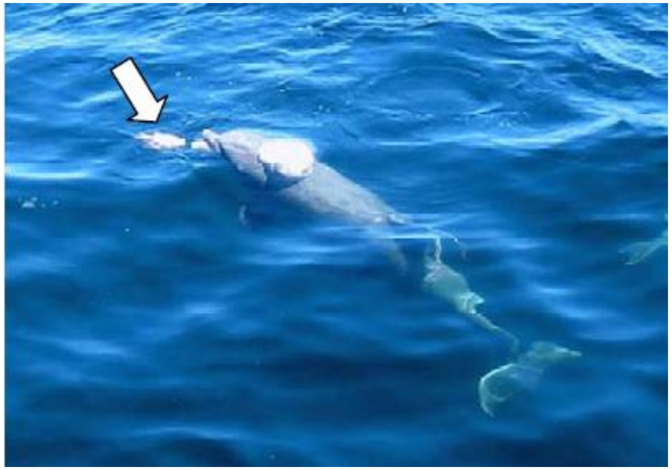
Reconstructing landings history: How far back do we go?

Incorporation in the stock assessment model

- Assume no error (exact match to catch)
  - Fewer parameters to estimate in search, improving model convergence and perhaps resulting in improved statistical consistency
  - Places high weight on uncertain data, detracting from model's ability to fit other data
- Statistical models for observation error (normal, lognormal)
  - More realistic accounting of uncertain data (e.g., some indices of abundance are actually better known than some segments of the catch)
  - Difficult to estimate the observation error variances because we don't have repeated measures. Usually input fixed variances derived from sample design.
  - More fiddling required (i.e., increased number of numerical issues that must be resolved, sometimes by reducing variances on catch)
- Statistical models for process error
  - Deviations from models of effective effort (we are just starting to dabble with this)

Discards ($\geq$ landings for many species)

- Multiple causes (bag limit, size limit, catch & release, IFQ)
- High discard mortality rate for some species owing to hook damage, barotrauma, and predation



- Highly uncertain
  - Recreational entirely self-reported (except for headboat observer program beginning in 2005).
  - Commercial self-reported from logbooks
  - Commercial reef fish observer data only available from the Gulf of Mexico after 2005, but coverage < 10%. Effectively no coverage in South Atlantic and Caribbean

Discards (≥ landings for many species)

Preprocessing

- Pooling conventions: observer data often must be pooled across regions and/or years (thereby giving a false sense of interannual variability)
- Representativeness of sparse observer coverage unclear and subject to changing motivations for discards under IFQ system (catch estimates based on observed CPUE can be >> total catch reported under the IFQ system)
- To reconstruct back in time (rescale self-reported?)

Incorporation in the stock assessment model

- ~~Assume no error (exact match to discards)~~
  - Places high weight on highly uncertain data, severely compromising model's ability to fit other data
- Statistical models for observation error (normal, lognormal)
  - What to assume about years prior to observer program
  - Super-year approach for sparse data (when you don't believe year-to-year trends)
  - Catch-free models, discard-implicit models (e.g., CATCHEM)

Indices of Abundance

Fishery-dependent CPUE Preprocessing

- Subsetting data to identify similar trips (excluding irrelevant trips)
  - species complex approach versus logistic regression to predict co-occurring species
- Standardizing to account for changes in catchability (typically use GLMMs)
  - Changes in targeting and operation of fleets (model as factors or split index?)
  - Effects of regulations (e.g., censoring model for bag limit)
  - Year/area interactions (random effects versus persistent trends)

Fishery-independent Survey Preprocessing

- Design-based estimates
  - Unbalanced in time and space, with important gaps
- Model-based estimates (standardizing to account for changes in catchability)
  - Typically use GLMs to standardize of environmental covariates (e.g., bottom temperature)

Indices of Abundance

Incorporation in the stock assessment model

- Statistical models for observation error
  - Normal
  - Lognormal
  - Bias-corrected Lognormal,
  - Gamma?
- Weighting schemes
  - Expert opinion (often equal)
  - Input variances (from GLM, or design-based)
  - Estimate variances (iterative re-weighting)
- Statistical models for process error in scaling coefficient q
  - Functional relationships
  - Penalized likelihood (e.g., random-walk)
  - Random-effects (Frequentist or Bayesian)

Length Composition

Preprocessing

- Presumed to representative (experiments in progress to test the extent this is so)
- Trip-based weighting?

Incorporation in the stock assessment model

- Statistical models for observation error
  - Multinomial? Other?
- Effective sample size
  - Number sampled (usually capped at maximum of 100-200)
  - Number of trips sampled (underestimate)
  - Estimated from model fits (reweighting schemes like McAllister et al. and Francis et al. often suggest effective sample sizes < 100)

Age Composition

Preprocessing

- Validation studies
  - Age at first ring formation, Frequency of ring formation
  - Quantification of reader imprecision and bias
- Evaluation of sampling design: is it representative of
  - Age frequency (trip-based weighting?)
  - Distribution of age at length (age-length keys)

Incorporation in the stock assessment model

- Statistical models for age frequency
  - Multinomial? Other?
  - Effective sample size (number sampled, number trips, estimated from model fits)
- Statistical models for age-length keys
  - Deterministic: Empirical age-length key used to convert length comp. to age comp.
  - Stochastic: Model fit to empirical age-length key

Shrimp bycatch

Preprocessing

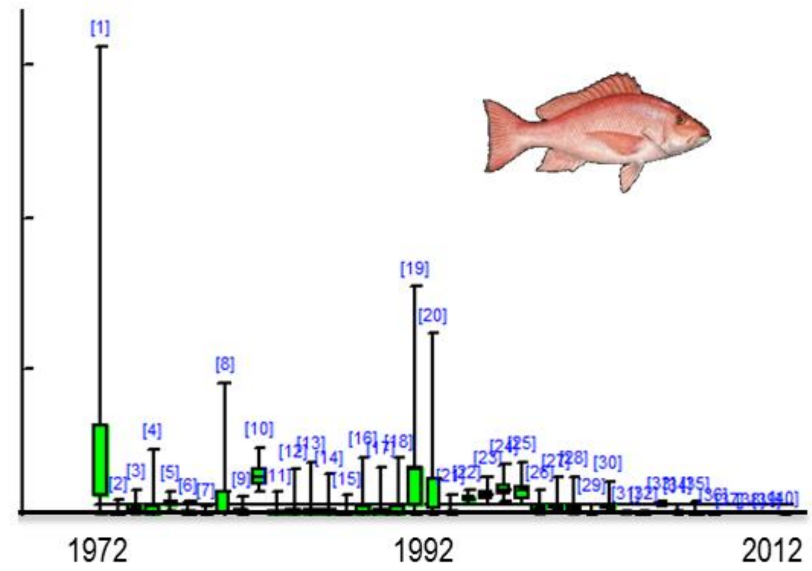
- CPUE by time/area strata (shrimp observer program, SEAMAP surveys)
- GLM or Bayesian equivalent used to fill gaps in CPUE
- Effort by time/area strata (shrimp catch / shrimp
- Bycatch = CPUE * Effort (by time/area strata)

Shrimp bycatch

Preprocessing

Incorporation in the stock assessment model

- Use entire time series of bycatch
    - Very high CV's for most species
    - annual trends not meaningful

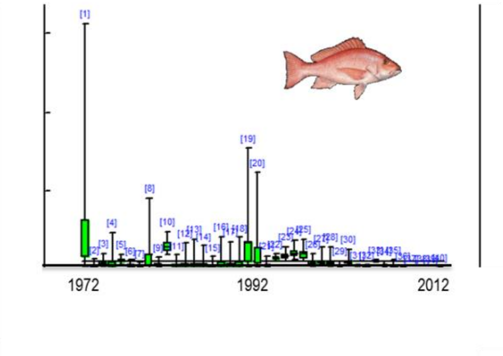


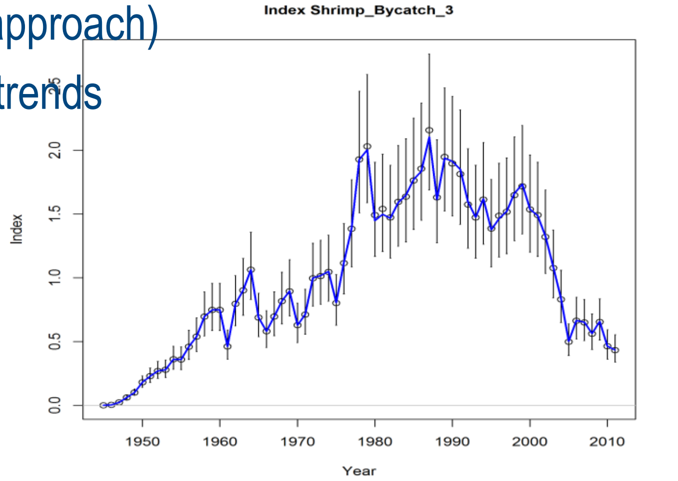
Shrimp bycatch

Preprocessing

Incorporation in the stock assessment model

- Use entire time series of bycatch
  - Very high CV's for most species
  - Annual trends not meaningful
- Use only mean level of bycatch over time series (reasonable CV for most species)
  - Model fits only the overall mean ("superyear" approach)
  - Effort data used to index fishing mortality rate trends

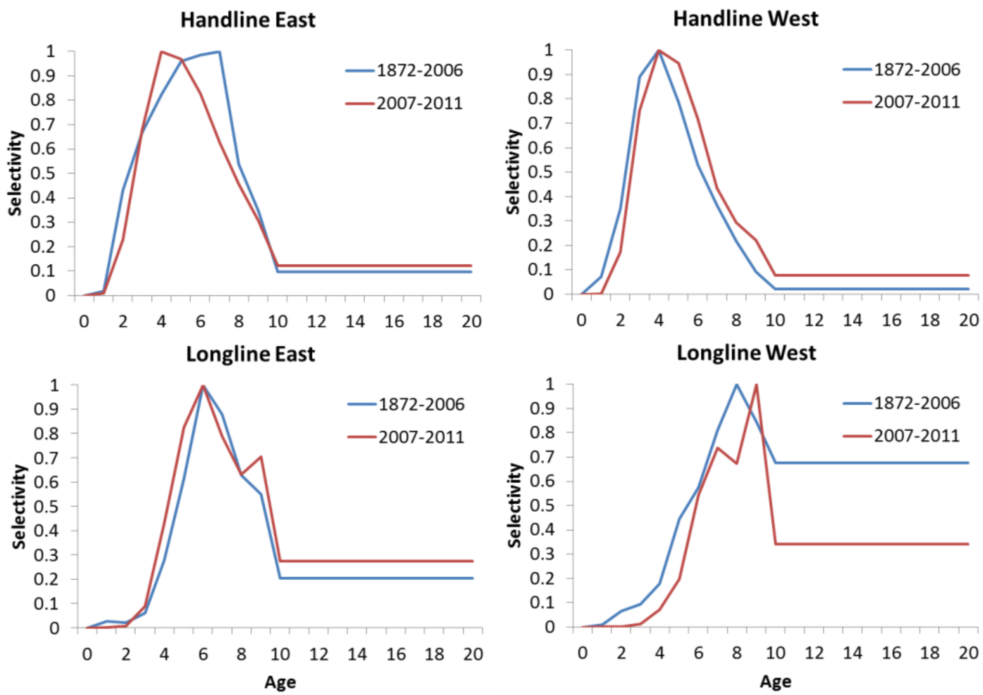


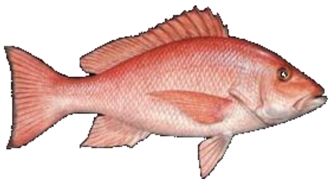


Assessment model structural assumptions

Selection

- Age-based versus size-based
- Age/size specific versus parametric
- Dome-shaped versus flat-topped; should we assume one until proven otherwise?



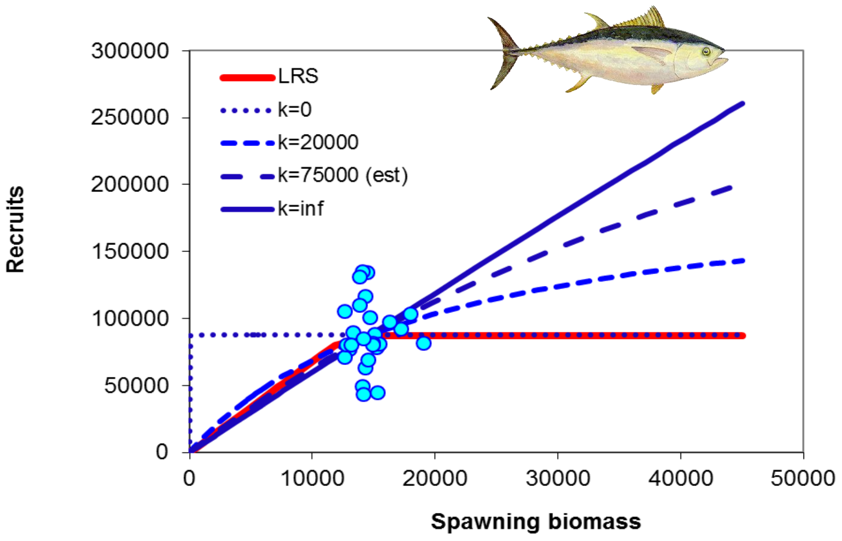


Assessment model structural assumptions

Selection

Recruitment

- Assume spawner-recruit curve?
  What functional form?



- Process error models
  1. Deterministic (no error) = ASPM
  2. Specify statistical distribution
     - normal or lognormal with bias correction
     - phased-in bias correction
     - gamma?
     - autocorrelation?
  3. Random effects (or Bayes equivalent)

Assessment model structural assumptions

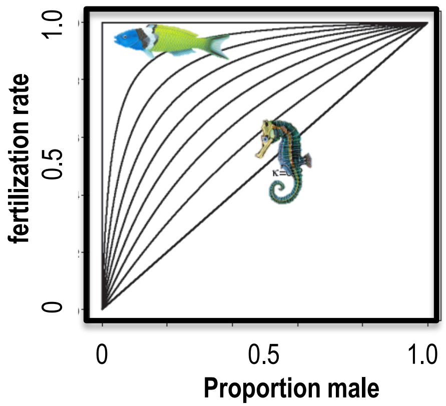
Selection

Recruitment

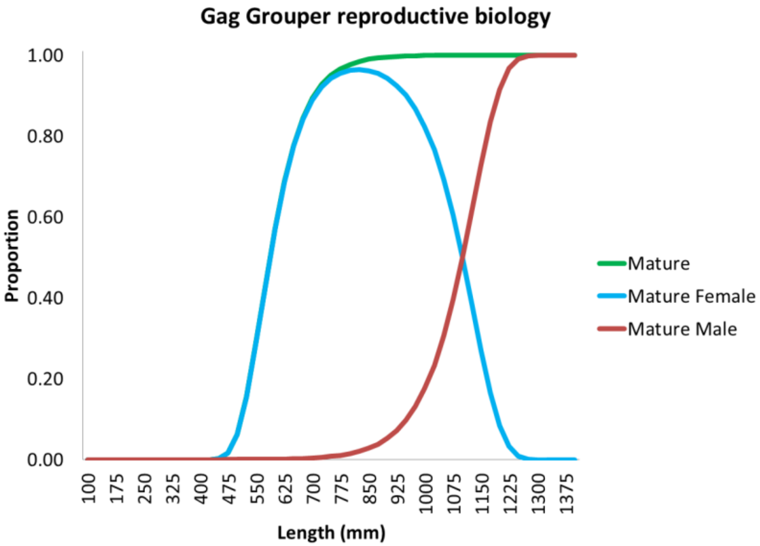
Reproduction/fecundity

Spawning stock biomass or spawning stock fecundity?

- Growth rate in egg production > growth rate in weight
- Males may also be limiting (fertilization success)



Protogynous hermaphrodites



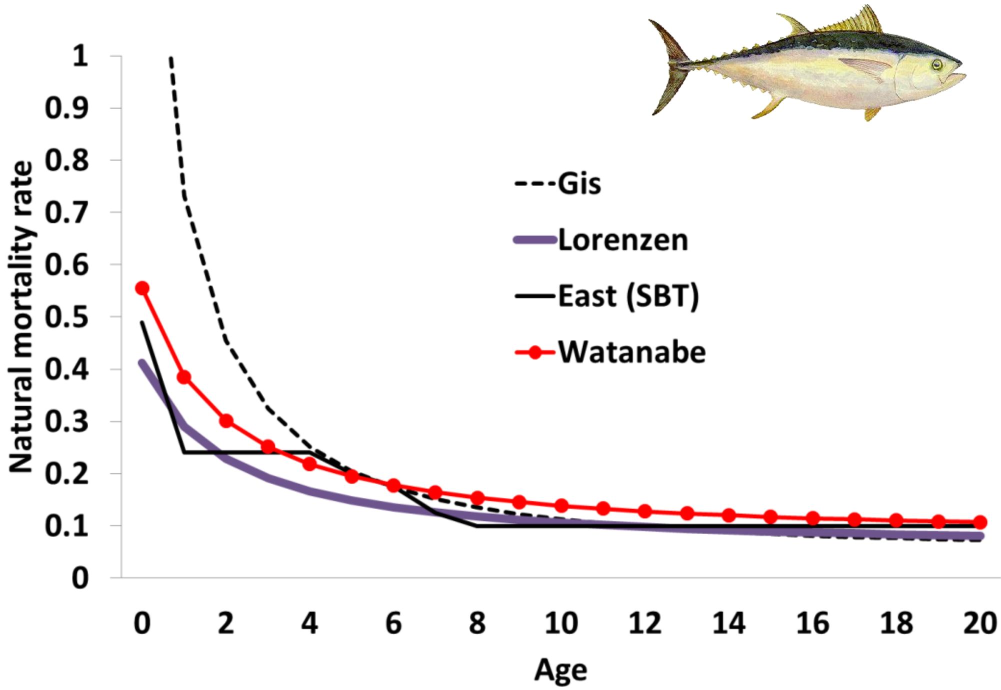
Assessment model structural assumptions

Selection

Recruitment

Reproduction

Natural Mortality rate



The Lorenzen curve is most often used in SEDAR assessments, but it is typically rescaled such that the average M on relevant age classes matches an “independent” estimate of M (folks believe the shape, but not so much the magnitude)

Assessment model structural assumptions

Selection

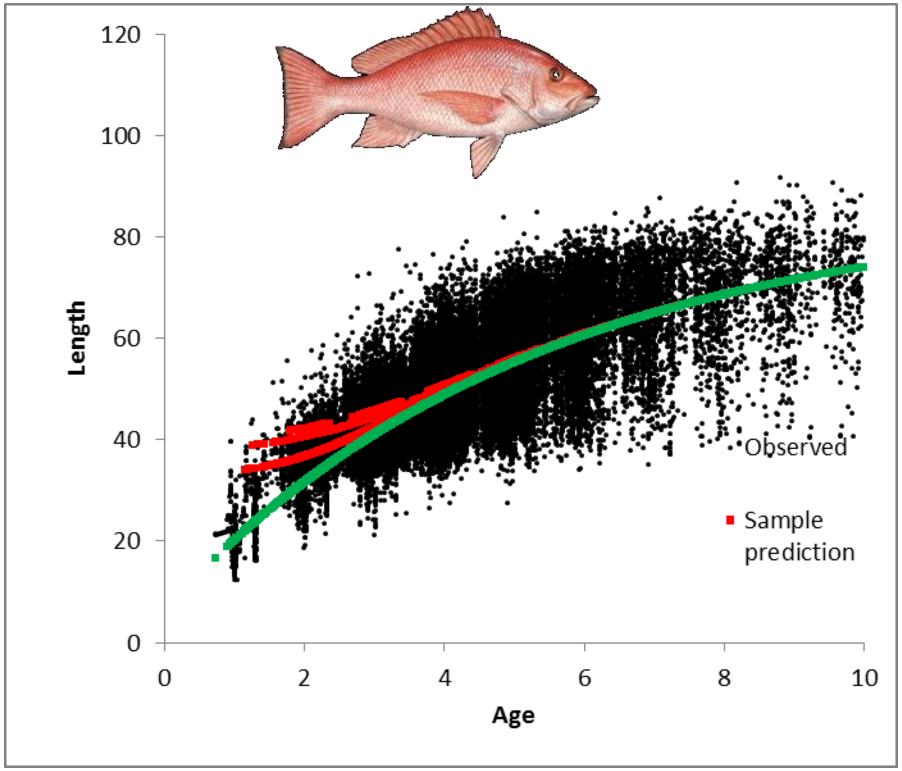
Recruitment

Reproduction

Natural Mortality rate

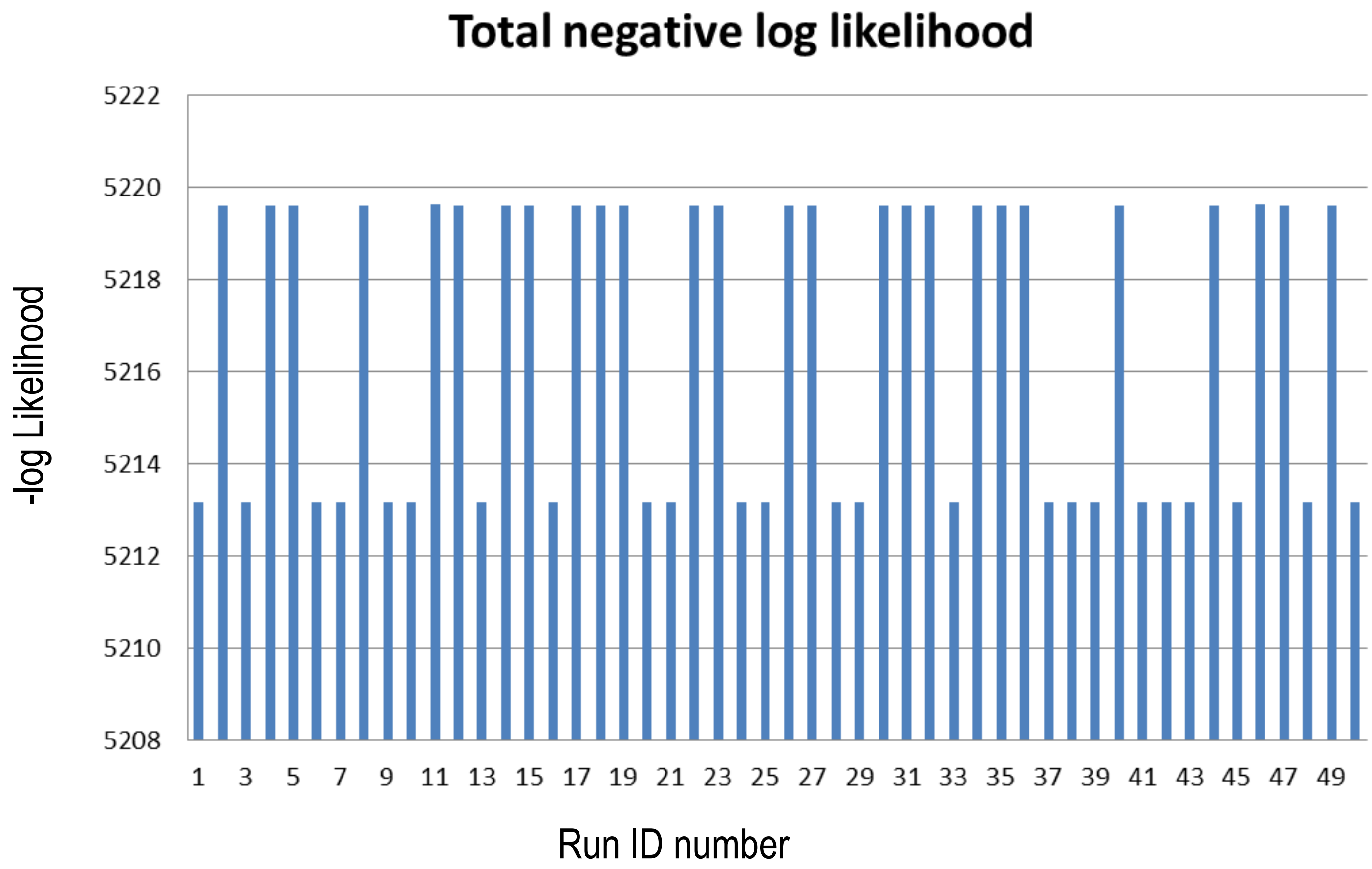
Growth: Von Bertalanffy

- much individual variation
- minimum size limits
- rapid early growth
  (linear interpolation to age 0)



Assessment model diagnostics

Jitter starting values (testing for global minimum)

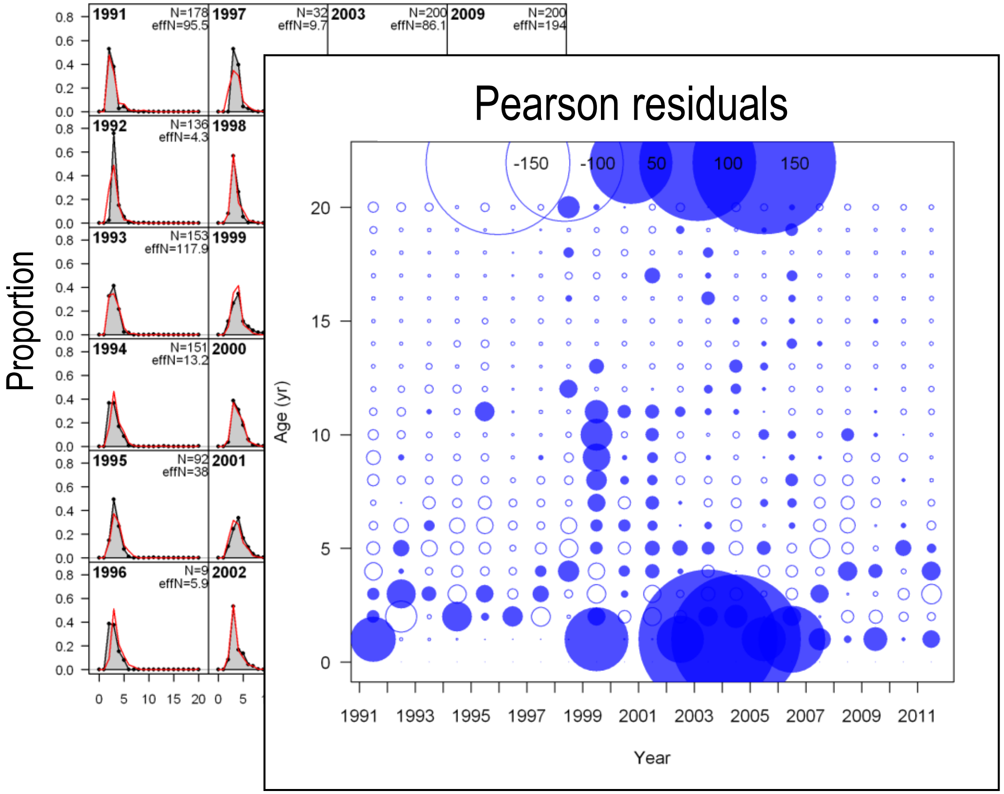


Assessment model diagnostics

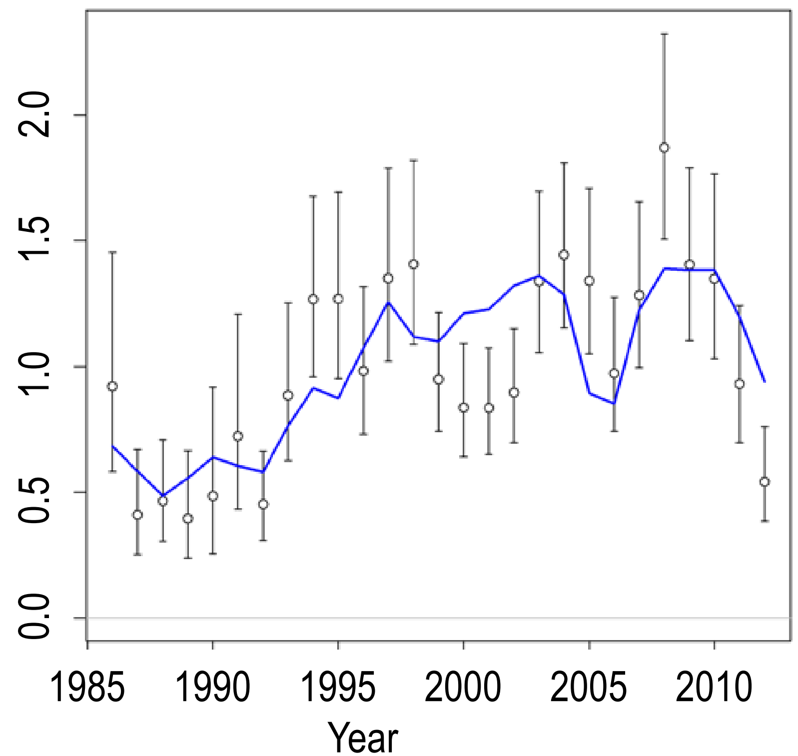
Jitter starting values (testing for global minimum)

Fits to data

Composition data



Indices

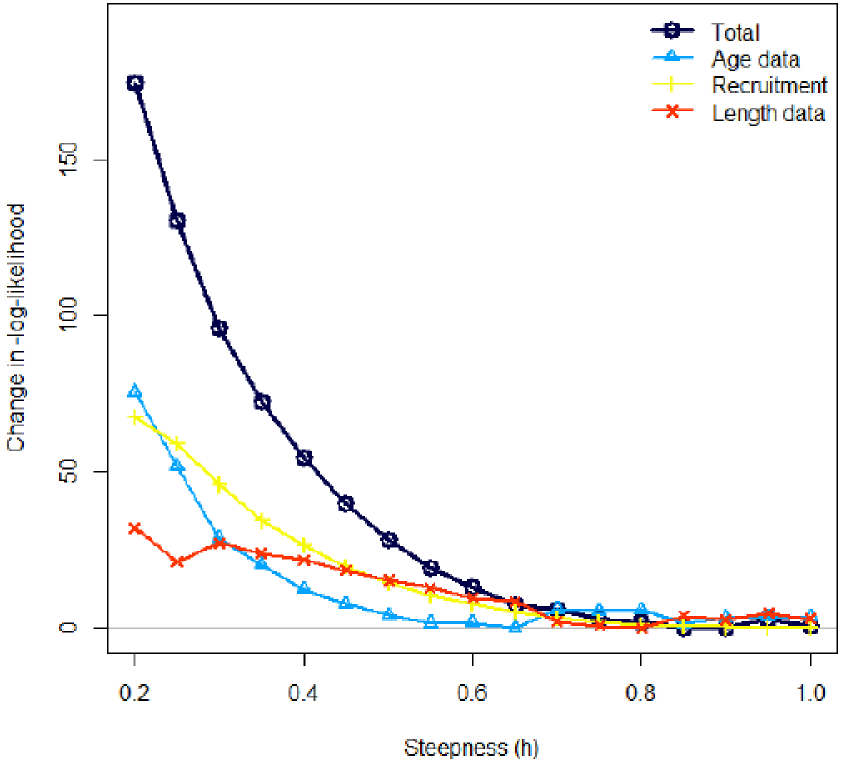


Assessment model diagnostics

Jitter starting values (testing for global minimum)

Fits to data

Likelihood profiling of key parms (steepness, R0)



Assessment model diagnostics

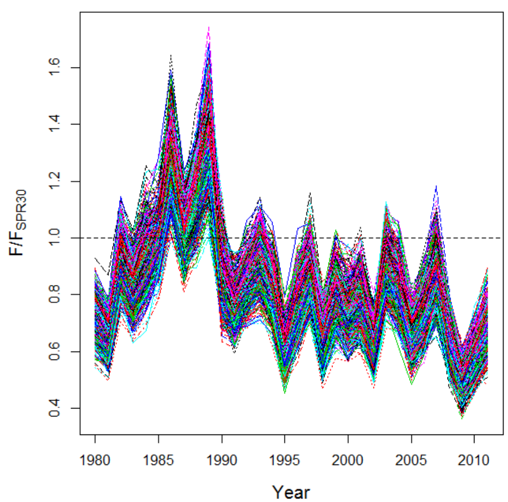
Jitter starting values (testing for global minimum)

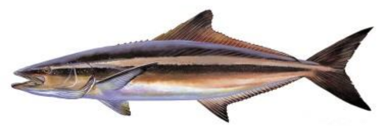
Fits to data

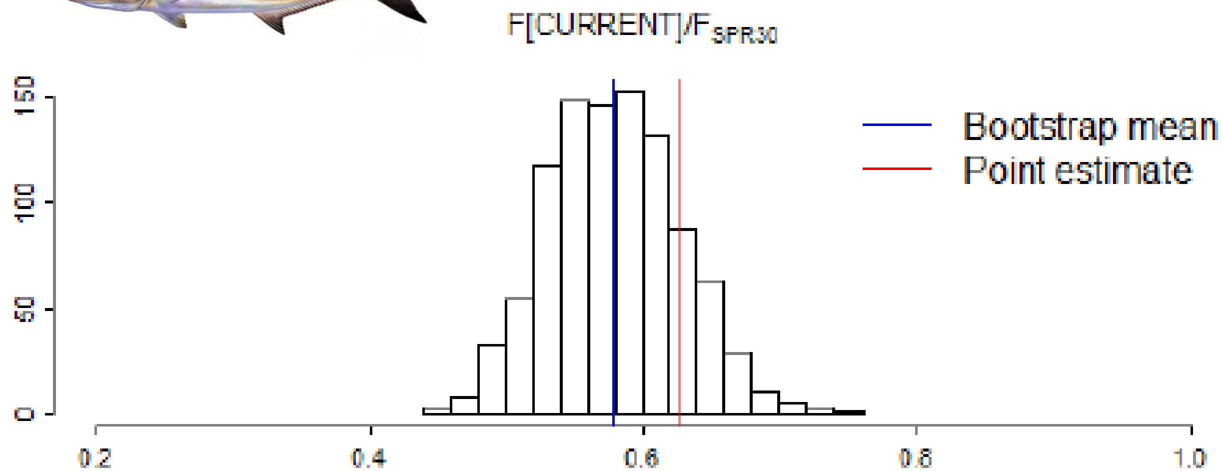
Likelihood profiling of key parms (steepness, R0)

Bootstrapping

Bias correct?







Assessment model diagnostics

Jitter starting values (testing for global minimum)

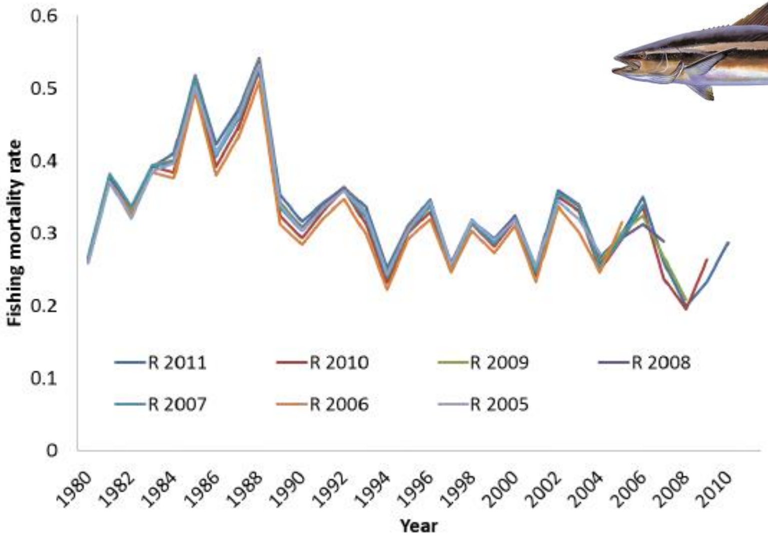
Fits to data

Likelihood profiling of key parms (steepness, R0)

Bootstrapping

Retrospective analyses

Bias correct?



Models for generating scientific advice

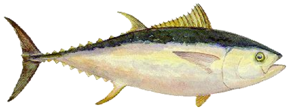
"Fishery management is an endless argument about how many fish there are in the sea, until all doubt has been removed – but so have all the fish." John Gulland

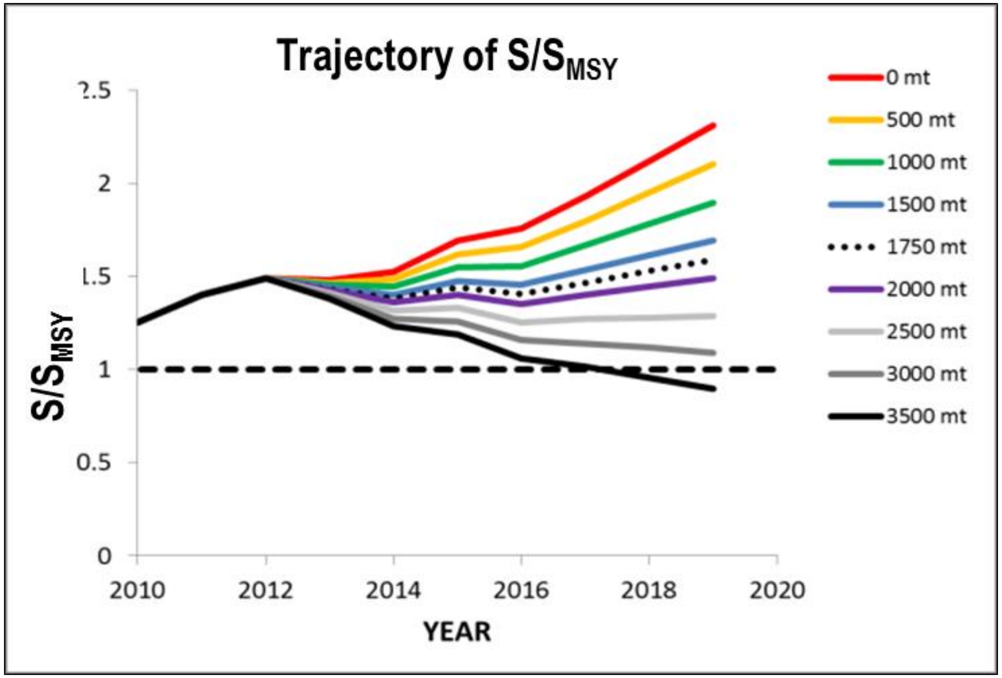
Projecting stock status into the future

Determining reference points

Characterizing uncertainty

Projections





Fishing models:

- Future selectivity = recent?
- Relative effort = recent?

(typically geometric mean of last several years)

Recruitment models:

- Recent levels
- Assume a particular spawner-recruit relationship

Stock status reference points: Setting the bar

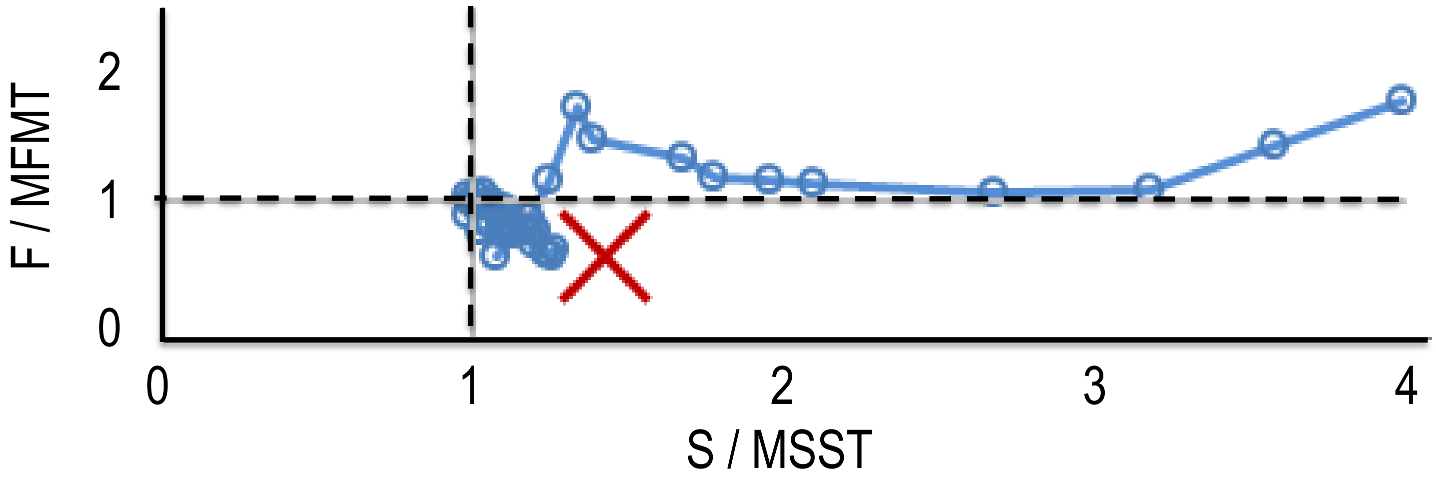
NS1 Guidelines Definitions:

MSY = maximum sustainable yield (MSRA goal)

MFMT = maximum fishing mortality threshhold, typically **F_{MSY}**

MSST = minimum stock size threshold, typically (1-M)***S_{MSY}** (S = spawning biomass)

OFL = overfishing limit (catch level when MFMT applied to current biomass)



Stock status reference points: Setting the bar

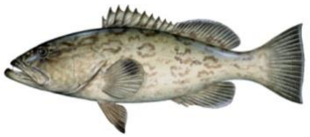
Determining MSY (F_{MSY} and S_{MSY}) :

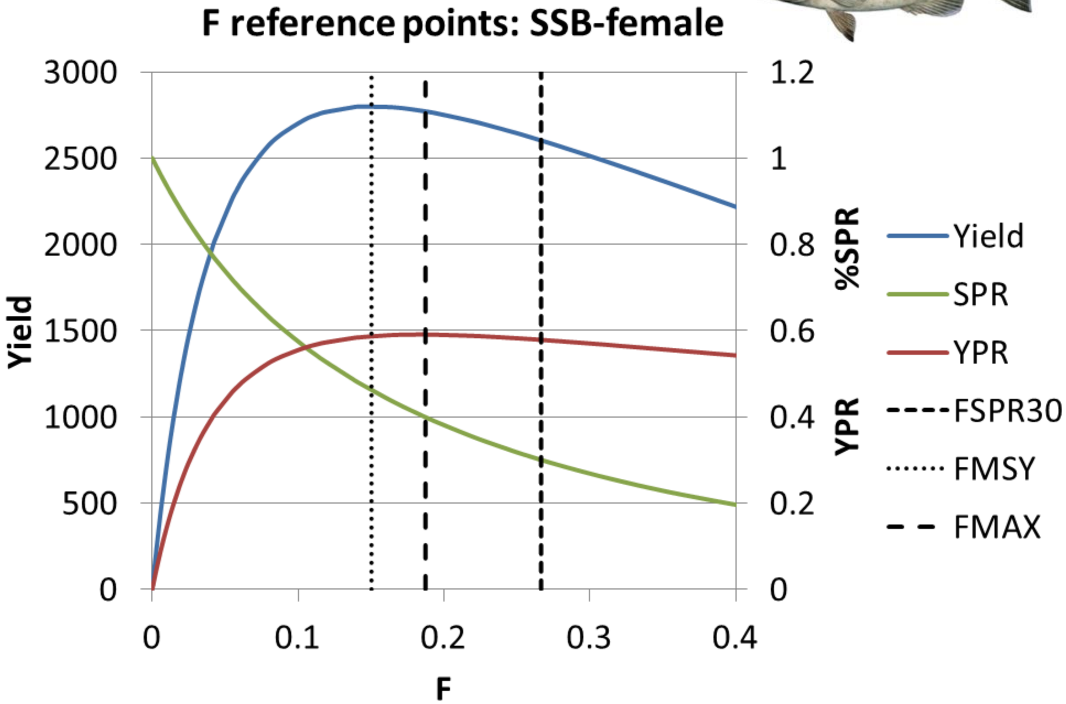
- Fishing models: Is the recent past the key to the future?
- Recruitment models: Is there a defensible spawner-recruit relationship?

  Yes: Compute MSY, F_{MSY} and S_{MSY}

  No: Proxy for MSY

  - Spawning potential ratio SPR
    - doesn't consider yield
    - typical values 20 – 40%
  - Maximum yield per recruit
    - doesn't consider spawners
    - $F_{MAX} = F_{MSY}$ (if recruitment is independent of spawners)

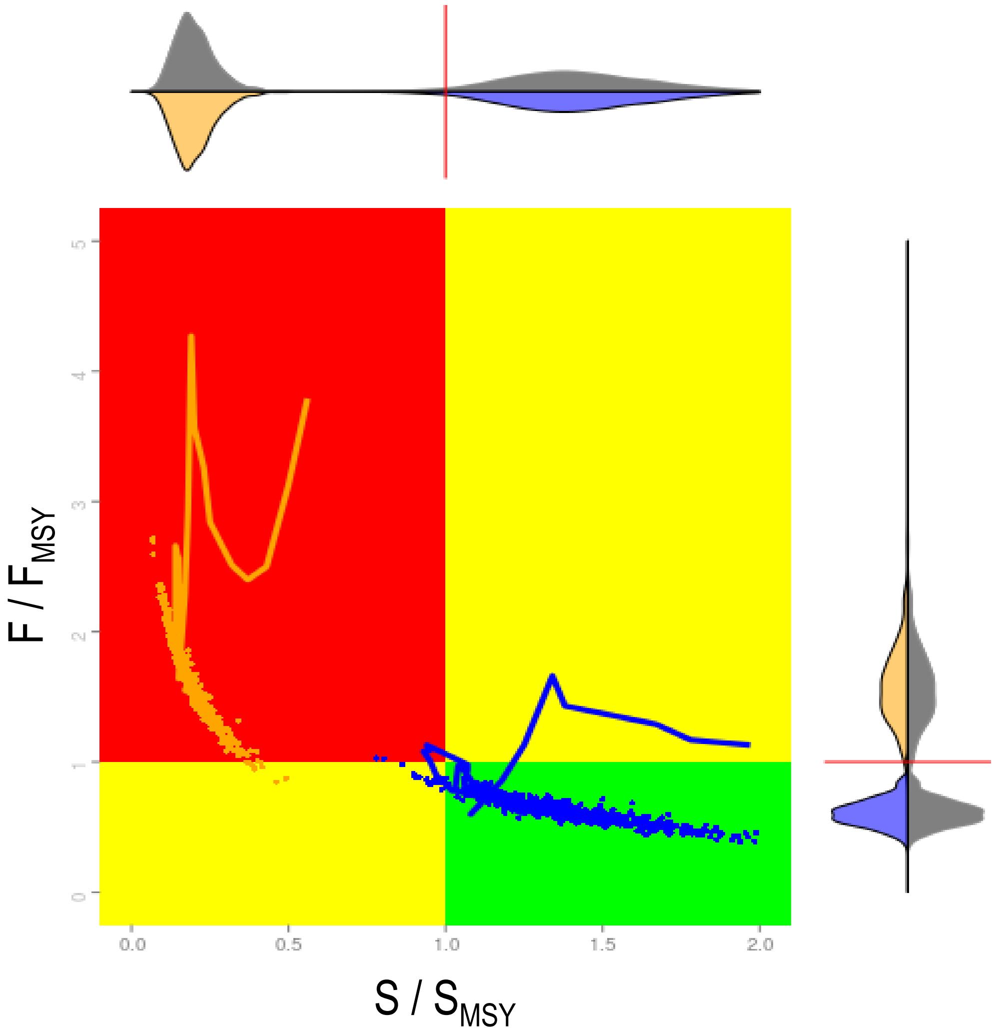




Quantification of uncertainty

Three basic approaches to estimate variance

- Single "best" model
- Multi-model inference
- Empirical

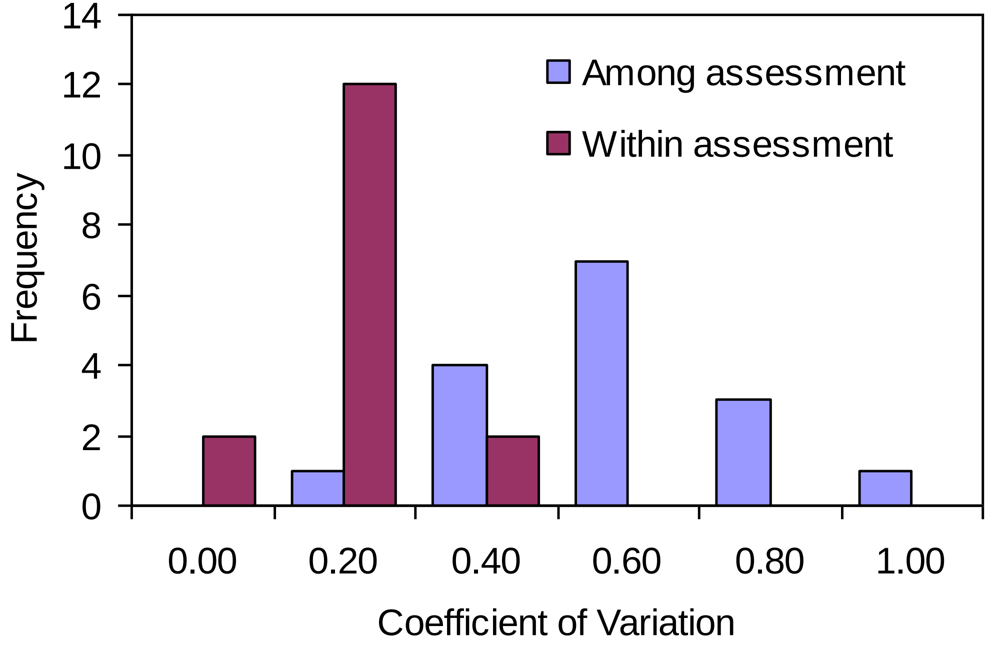


Quantification of uncertainty

- Single "best" model
  - Inverse-Hessian
  - Bootstrap
  - MCMC

Gives only the "within assessment" variance, assuming model is correctly specified

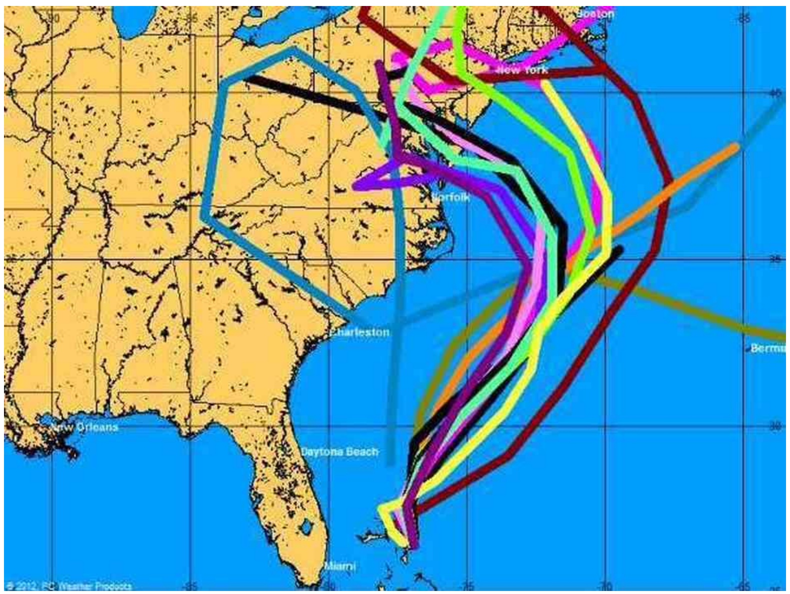
In practice, there tends to be more variation among alternative models than within models



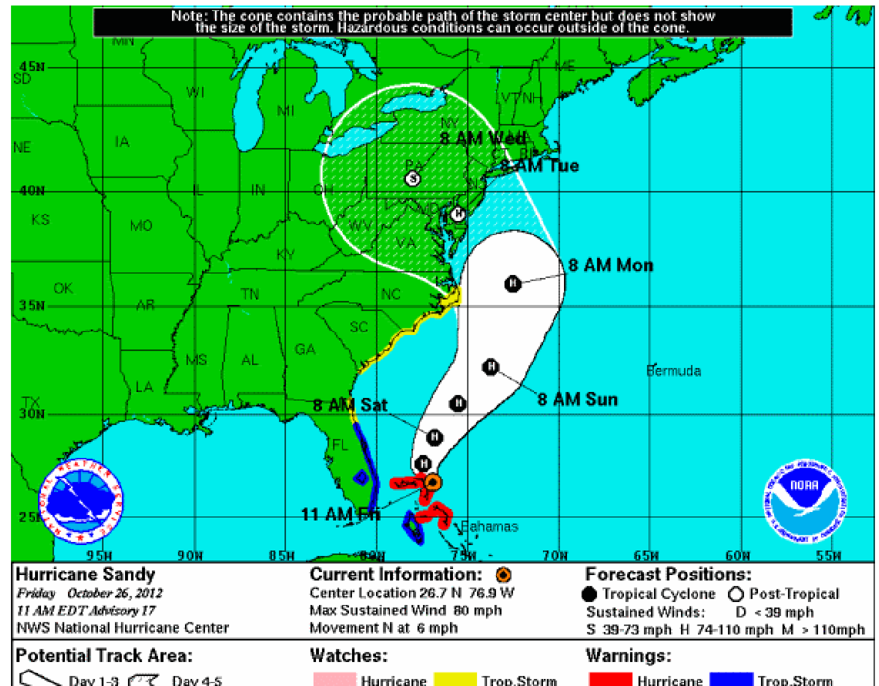
Quantification of uncertainty

- Single “best” model
- Multi-model inference (ex. hurricane forecasts)

How can we weight individual tracks



to get this?



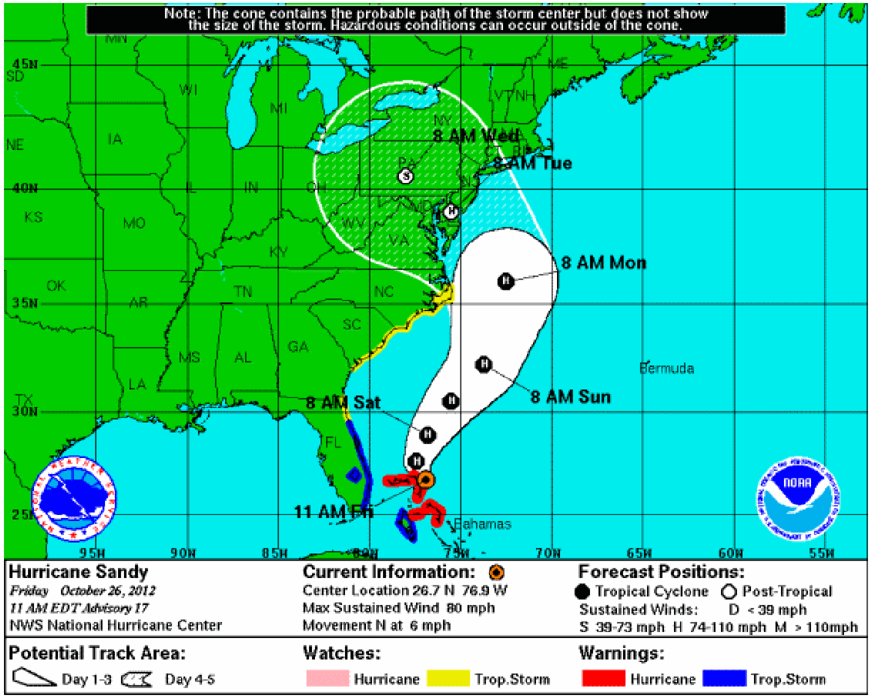
Quantification of uncertainty

- Single “best” model
- Multi-model inference

It’s not hurricane science.

It’s harder!

We know the trajectory of the hurricane,
we don’t know the trajectory of the stock

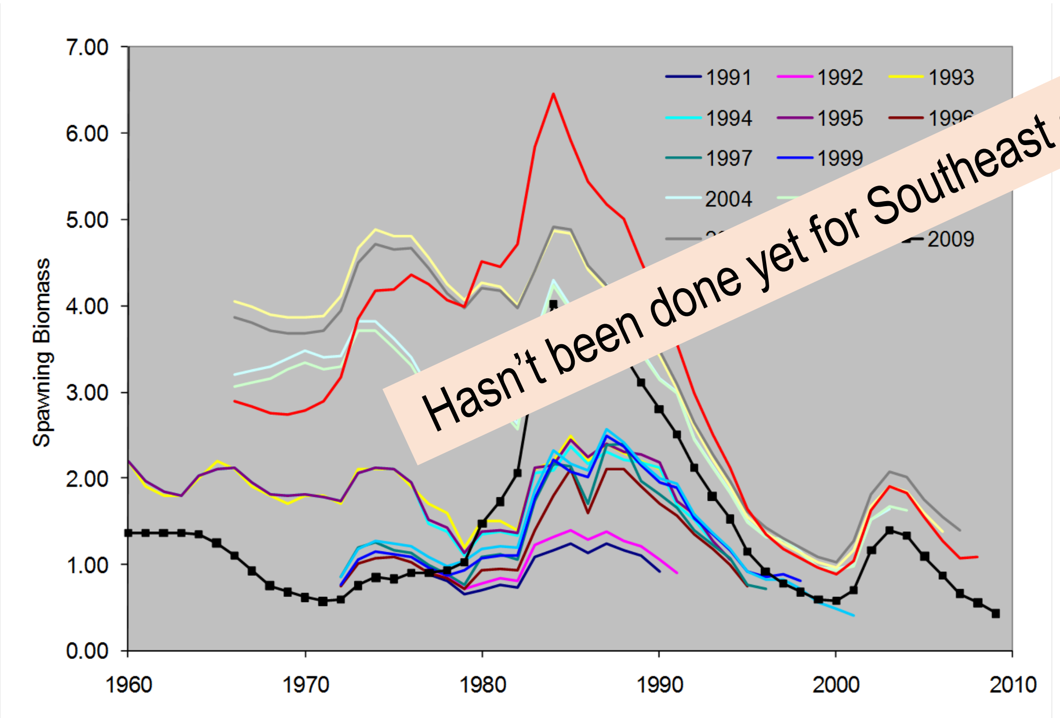


Quantification of uncertainty

- Single “best” model
- Multi-model inference
  - Information criteria (e.g., AIC): data must be the same
  - Bayes factors: priors can be highly influential
  - Expert opinion

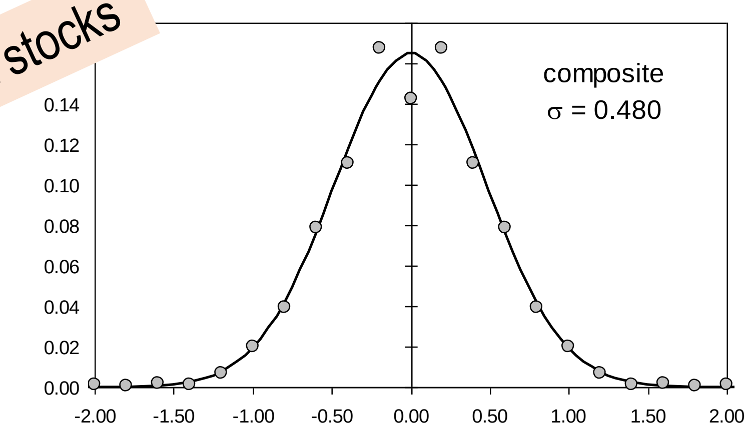
Quantification of uncertainty

- Single “best” model
- Multi-model inference
- Empirical (Pacific Fishery Management Council, Ralston et al.)



Courtesy Steve Ralston

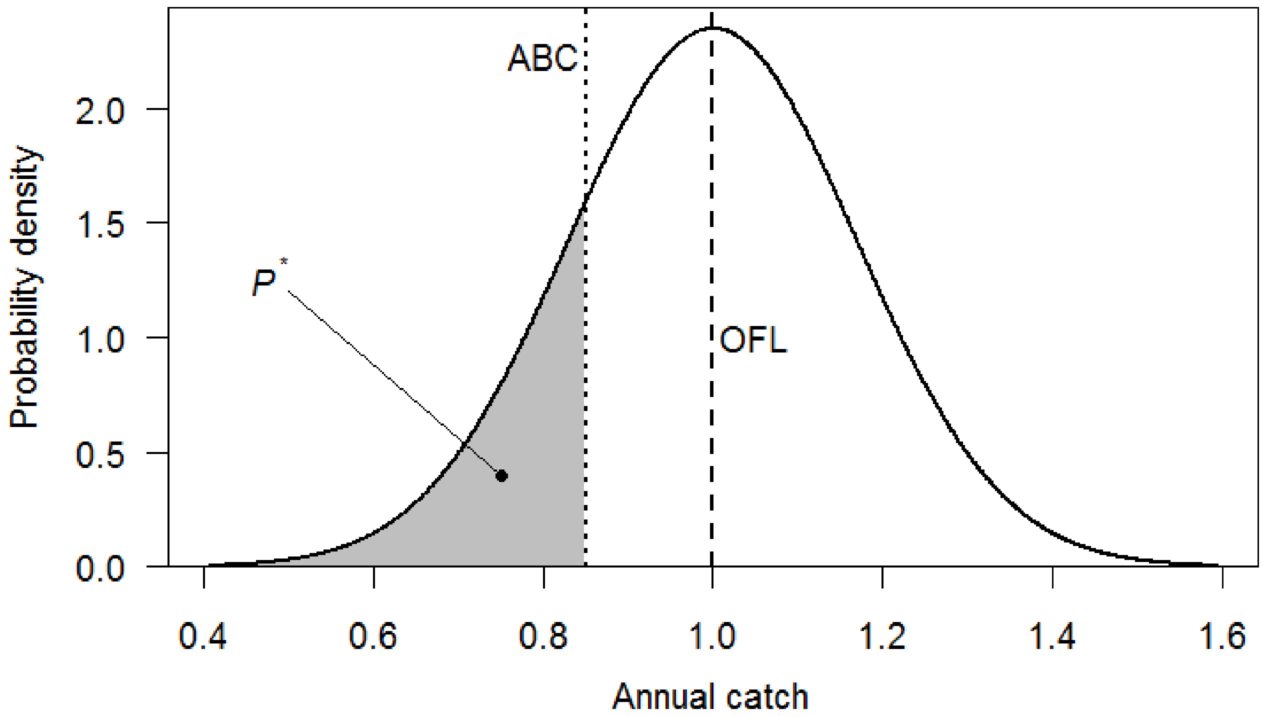
Hasn’t been done yet for Southeast stocks



Many species and many common years

Translating uncertainty into ABC Advice

The PSTAR approach



Translating uncertainty into ABC Advice

The PSTAR approach

-currently codified in ABC control rules

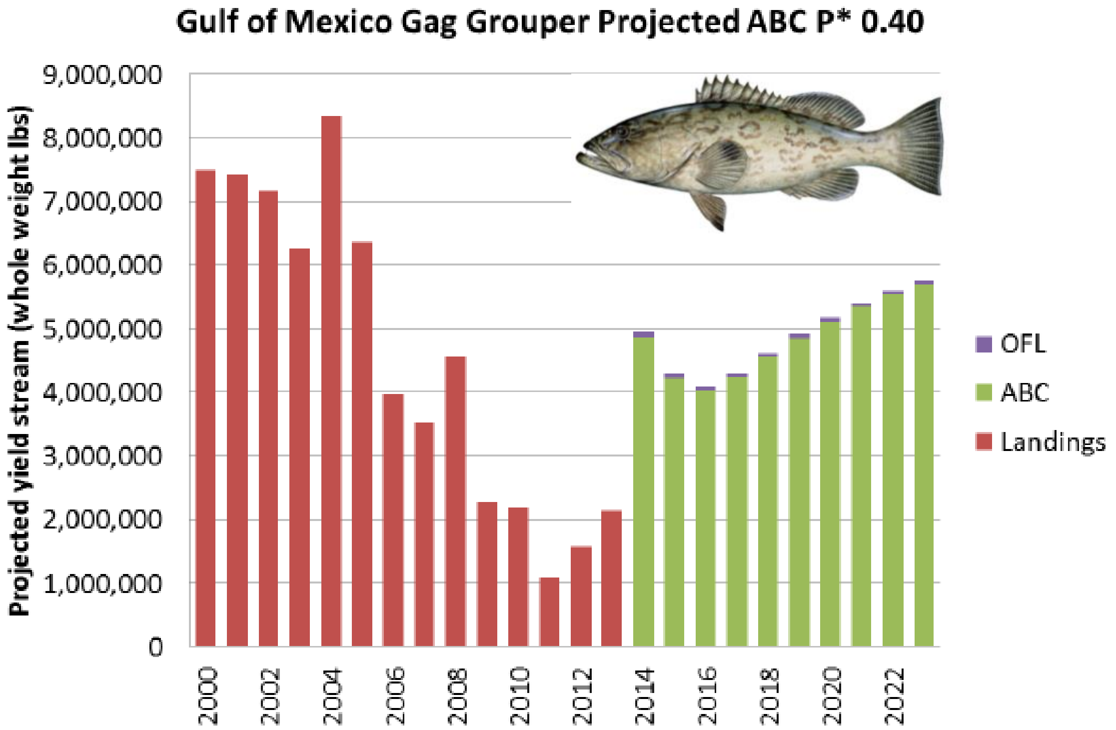
-explicit accounting of scientific uncertainty

-much work

-variances uncertain

-variances often small

(small buffer)



Translating uncertainty into ABC Advice

Should we PSTAR or should we KISS?

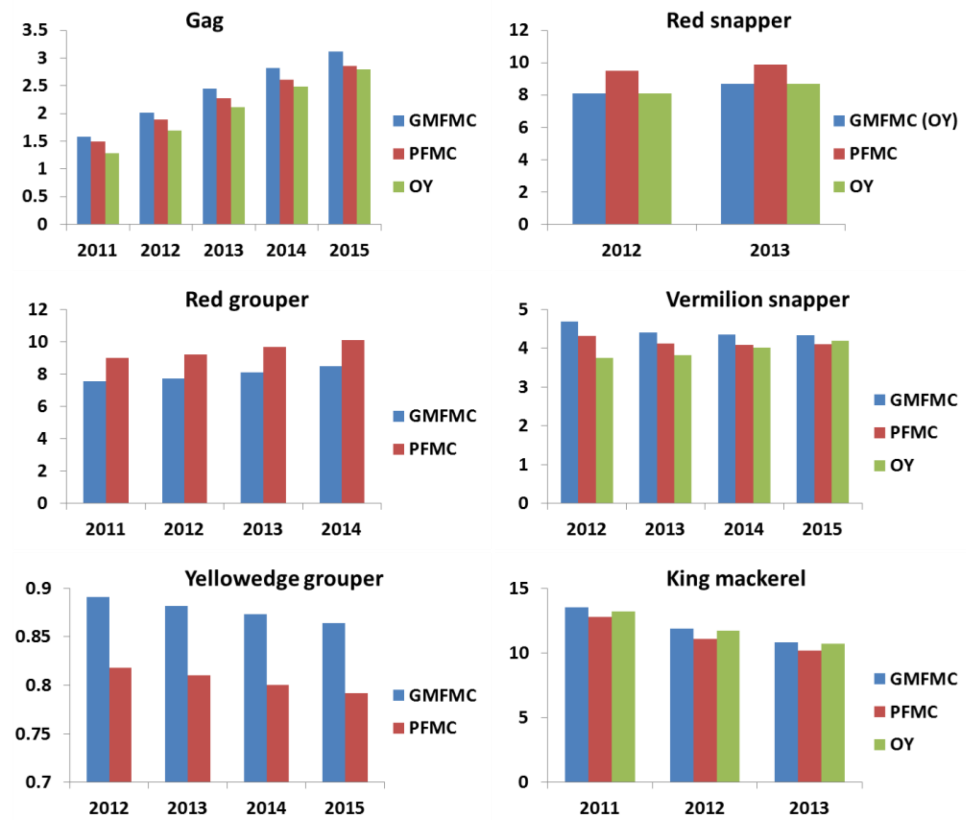
PSTAR

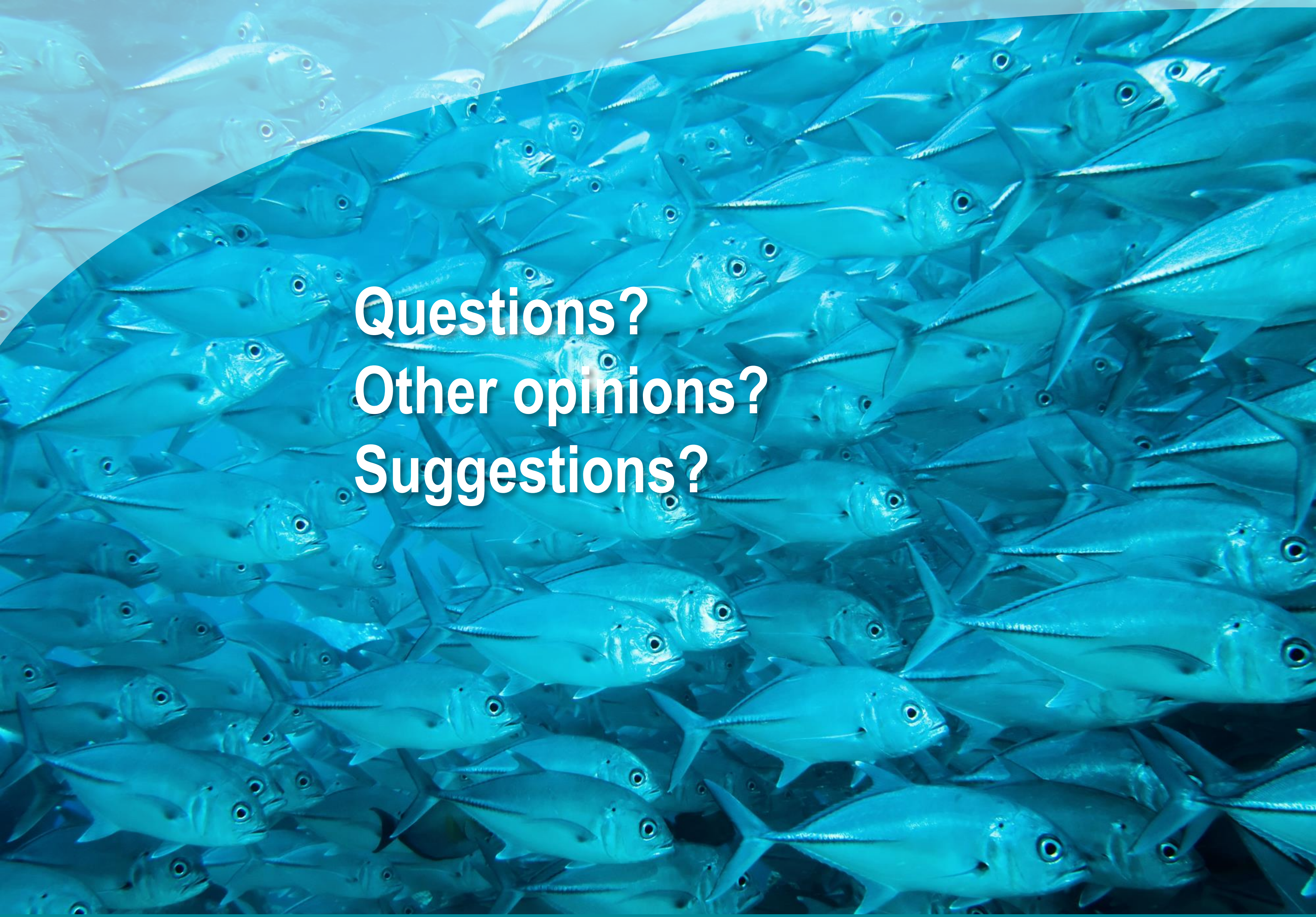
- current, explicit
- much work
- variances uncertain, often small

KISS=robust Harvest Control Rule

e.g., $F_{target} = 0.75\ F_{limit}$

- less work
- familiar to Councils as OY
- gives comparable (often lower) ABC
- not adequately simulation tested





Questions?
Other opinions?
Suggestions?